Research on the Application of Electric Vehicle Recognition in Lifts Based on Yolov8 Model

Xuanjie Xu [1, #], Shuxiang Yang [2, #, *], Haoran Shu [3, #]

[1] College of Electronics and Information Engineering, Shenzhen University, Shenzhen, China, 518060

[2] Faculty of computational mathematics and cybernetics, Shenzhen MSU- BIT University, Shenzhen, China, 518172

[3] School of Mathematical and Physical Sciences, Chongqing University of Science and Technology, Chongqing, China, 401331

* Corresponding Author Email: rhiya_3505@163.com

Abstract. The behaviour of electric bicycle into the lift is a great security risk, but also one of the causes of many fires in the building, bringing a serious threat to people's life and property safety. In this paper, we use YOLOv8 model to identify and detect e-bikes in the lift scene. Compared to YOLOv5, YOLOv8 improves performance and employs an anchor-free mechanism, reducing hyperparameter tuning complexity and enhancing small object detection in elevator scenarios. The experimental test results show that: the mAP@50 reaches 87%, and the operating speed reaches 125 FPS, which meets the deployment requirements of the actual monitoring system. This study provides a new solution in the field of lift safety monitoring, which can significantly improve the efficiency of detecting e-bikes in lifts and prevent safety accidents such as fires caused by e-bikes.

Keywords: YOLOv8; Electric bicycle; Lift safety; Risky behaviour detection.

1. Introduction

With the acceleration of urbanisation, the electric vehicle (EV) industry has entered a mature stage, and electric vehicles have become one of the most important means of transport for residents travelling short distances. Since the common lift space is relatively small and closed, once the e-bike fires, the rapid spread of high temperature and toxic gases will lead to people inside the lift having nowhere to escape. In order to comprehensively manage the fire hazards of electric bicycles, the General Office of the State Council issued the ‘Electric Bicycle Safety Hazards Full Chain Remediation Action Programme’ in April 2024, focusing on the construction of standards and systems, the management of production and sales, parking and charging norms, as well as scrapping and recycling processes to carry out a full chain of safety management. By November 2024, the action had begun to bear fruit, with the number of fires per million e-bikes dropping from 5.2 to 3.4 nationwide. Based on this, in order to respond to the national policy and further reduce the risk of e-bike fires, monitoring e-bike into lift events using computer vision target detection algorithms has been regarded as a proven technological solution. The aim of this study is to develop and optimise a computer vision-based safety monitoring method for electric vehicles in order to enhance the safety of the environment in which they are used.

At present, domestic and international research on elevator access control systems has roughly gone through three development stages [1]: mechanical device blocking, sensor detection, and artificial intelligence detection based on computer vision. Mechanical device blocking: prevent electric vehicles from entering the elevator by installing a metal guardrail in the access channel outside the elevator floor door. This method is inexpensive and easy to install, but there are obvious limitations, such as low efficiency, easy cracking, and may affect the normal use of the elevator by special groups (e.g., wheelchair users). Sensor Detection: With the help of sensors installed inside the elevator, the shape and contour of objects entering the car are detected. This method improves automation to a certain extent, but the detection accuracy is low and prone to false alarms and omissions. Artificial intelligence detection based on computer vision: images or videos are captured by cameras installed on top of the elevator car, and image processing and analysis techniques are used to achieve accurate identification of electric vehicles. This method relies on deep learning technology, which significantly improves the accuracy and real-time performance of detection and has become a current research hotspot. In recent years, the wide application of deep learning technology in the field of target detection has significantly improved the accuracy and real-time performance of detection. In the field of public safety, algorithms such as the improved SSD model and EfficientNetV2-SSD have realized efficient detection in scenes such as mask wearing and helmet monitoring. In the field of electric vehicle entry elevator detection, researchers have proposed a variety of schemes based on image recognition, sensor fusion and deep learning. Common techniques include YOLO series algorithm, neural network fusion, etc. Existing research mainly focuses on improving the detection accuracy and real-time performance, such as OR-CNN method improves the detection accuracy in occluded environments. Among them, the research based on YOLO series algorithms is particularly outstanding, and the various improved algorithms of YOLOv8 have demonstrated excellent performance in small target detection and complex scenes in fabric defects, foggy targets, infrared ship detection, and other detection applications.

Based on the high efficiency and excellent real-time performance of the YOLOv8 model in the target detection task, this paper will use YOLOv8 to carry out research on the detection of electric vehicles entering lifts. It is expected to improve the real-time response capability and detection stability of the embedded system in complex background and narrow space while ensuring the detection accuracy, so as to reduce the possibility of safety accidents. The aim of this study is to explore efficient algorithms suitable for electric vehicles into lift detection, especially on complex background and occluded object detection, and to further optimise the performance of the YOLOv8 model in such real-world scenarios. The significance of the study is to improve the safety and efficiency of lift management and provide technical support for automated detection of similar scenarios.

2. Literature review

In the field of safety detection, the introduction of deep learning technology has become a key factor in improving the accuracy and real-time performance of target detection, which is widely used in the fields of mask wearing detection and helmet wearing monitoring. The improved SSD model proposed by Li Yuyang [2] in 2022 uses ResNet50 to replace the original VGG-16 feature extraction network, which improves the depth and robustness of feature extraction, while the introduction of the BAFF module enhances the detection expressiveness of the mid-level feature maps, which improves the accuracy and efficiency of mask wearing detection in high-density public places. However, the multi-layer feature fusion of ResNet50 and BAFF module increases the computational complexity and is not suitable for application in resource-limited devices such as embedded. This drawback is improved in the study of improved helmet detection algorithm based on EfficientNetV2-SSD proposed by Xin Wang and Yinan Feng [3] in 2022, which adopts EfficientNetV2 to replace the original VGG-16 feature extraction network of SSD to reduce the network parameters, and introduces the FPN pyramid structure for feature fusion, which improves the accuracy and efficiency of small targets and dense scenes and increased the detection speed, but its adaptability is dependent on the dataset, which may require re-adjustment of parameters in other detection tasks, increasing unnecessary workload.

In the research field of electric vehicle access to lifts, several papers in recent years have proposed different detection and prevention schemes, Hongcheng Zhao [4] et al. proposed a multifunctional alarm system based on STM32 microcontroller in 2023, which was designed to combine the functions of image recognition, voice alarms, and lift operation restrictions. This approach improves the efficiency of lift management compared to traditional simple alarm systems; however, the system relies on the accuracy of the image recognition module and the detection success rate may fluctuate in different scenarios. Ziyu Ma [5] in 2021 combined multiple neural networks and improved D-S evidence theory to improve the robustness of the electric vehicle detection model in lifts by fusing sound, acceleration, and image data, however, the high dependence on sensors and model complexity makes it challenging to deploy. Lin Chen [6] in 2022 simplified YOLOv4 and achieved real-time detection of electric vehicles in lifts via YOLOv4-tiny, with fewer model training parameters to optimise the detection speed, but it is still deficient for small objects and multi-target collapsed detection. He Bigan [7] et al. proposed a restricted access system for electric vehicles in lift car based on YOLOv5 target detection algorithm and Jetson Nano embedded development board in 2023, which made great progress in improving detection accuracy and intelligent design, but the quality of the dataset and hardware dependence are still its potential bottlenecks. The OR-CNN method proposed by Lv Qiaorun [8] in 2024 integrates the a priori structural information of the detection target with the visibility prediction into the Fast R-CNN module in the detector, and the result is that the OR-CNN network is more adaptable to the narrow and easy-to-obstruct target environment such as lifts, as compared to the recognition accuracy of YOLOv5. Shuguang Zhang [9] in 2024 proposed an improved MobileNet-SSD method, through CycleGAN data enhancement, BOHB hyperparameter optimisation and LReLU activation function, successfully improved the accuracy and real-time performance of electric vehicle recognition, especially suitable for embedded devices, but its optimisation process is more complex and hardware-dependent.

In the research of electric vehicle entry lift detection, YOLO algorithm excels in detection results due to its efficient target detection capability and strong real-time performance, the latest research progress of YOLO algorithm is the application of YOLOv8, Yuna Liu and Shuangbao Ma [10] proposed the lightweight YOLOv8n in 2024, which is improved by the C2fGhost module, the SimAM attention mechanism and Lightweight Shared Convolutional Detection Head (LSCDH), which greatly reduces the computational complexity and improves the detection of small targets without significant performance degradation, and this improvement allows YOLOv8n to perform well in resource-constrained scenarios such as fabric blemish detection. Zhen Liu [11] et al. in 2024 proposed a foggy target detection method based on YOLOv8s, which improved target detection in foggy environments by introducing the Edge-Dehaze defogging module and the Hybrid Attentional Feature Fusion Module (HAFM) to effectively deal with foggy weather and occlusion problems. Similarly, Haiqun Wang [12] improved the YOLOv8 model for infrared ship detection in 2024 by improving the YOLOv8 model, and the improvements included the introduction of the MCA attention mechanism in the YOLOv8 backbone network and the use of BiFPN instead of the original PAN structure, which significantly improved the detection performance for small targets and complex backgrounds. Ma Yaoming [13] in 2024 effectively improved the recognition of smoke flames in multi-context scenarios based on YOLOv8s by means of the C2fFR lightweight feature extraction module and the MCFM multi-scale context fusion module. Bo-Yuan Wang [14] improved the YOLOv8s model in 2024 by proposing the fine-grained convolution (SPD-Conv) and MPDIoU loss function, which improves the detection accuracy of the surface personnel and seepage of the earth-rock dams and adapts to the special scenarios captured by high-altitude unmanned aerial vehicles (UAVs).

3. Modelling

3.1. Introduction to the model

YOLOv8s is a lightweight, high-precision target detection model for real-time scenarios. Its one-stage structure supports global analysis of lift surveillance video to quickly identify electric vehicles, lifts and other targets. In complex environments inside lifts, YOLOv8s efficiently handles multi-target mixing and lighting changes, providing accurate support for real-time detection of electric vehicle violations.

3.2. Model structural design

In this paper, the classical YOLOv8 is used as the detection model, and its network structure can be divided into three main parts: the backbone network (Backbone), the feature fusion module (Neck) and the detection head (Head).

YOLOv8 improves on its predecessor model in many ways by introducing the Cross Stage Partial Connections with Fusion module(C2f) in the Backbone network and Neck structure instead of the Cross Stage Partial module (CSP) in YOLOv5. In the detection head part, YOLOv8 adopts the Anchor-Free mechanism for bounding box prediction, which does not rely on the predefined Anchor box, thus simplifying the design and improving the accuracy of small target detection. In the loss function part, YOLOv8's loss calculation includes two parts, the classification branch and the regression branch, and the classification branch adopts the BCE Loss, while the regression branch adopts the Distribution Loss. The classification branch uses BCE Loss, while the regression branch uses 'Distribution Focal Loss 'loss function and CIoU Loss.

3.3. Model structure details

Backbone is the core part of feature extraction in YOLOv8, which is responsible for the gradual extraction of low-level and high-level semantic features from the input image. Backbone takes the stacked Convolution-BatchNorm-SiLU modules (CBS) and C2f modules as the main structural units, among which the CBS module consists of a convolutional layer, a batch normalisation layer and a SiLU activation function, which is able to extract the image features efficiently.The C2f module achieves cross-stage feature fusion by means of feature segmentation (Split), stacked Bottleneck processing, and final feature splicing (Concat), which improves the feature expressiveness and information mobility of the network.

Neck is an important bridge between the Backbone and the detection head. The Neck in YOLOv8 adopts a C2f-based design to further optimise the features extracted from the Backbone through efficient feature enhancement. Neck combines high-level semantic information with low-level details through up-sampling and Concat. This multi-scale fusion mechanism ensures that the model is robust to the task of detecting large targets in electric vehicles.

The head module is responsible for generating the final target detection results, including category prediction, bounding box regression, and confidence estimation. YOLOv8's Head adopts a multi-scale prediction design, outputting detection results from feature maps at different resolutions (e.g., 80×80 and 20×20). The feature maps at each scale are first optimised by the C2f module, and then the categories, bounding boxes and confidence scores of the targets are generated by the Detect module. The YOLOv8 the network structure is shown in Figure 1. Where the Spatial Pyramid Pooling-Fast (SPPF) module enhances multi-scale feature fusion. Classification Loss (Cls. Loss) optimizes category prediction accuracy. Bounding Box Loss (Bbox. Loss) improves localization by minimizing prediction errors.

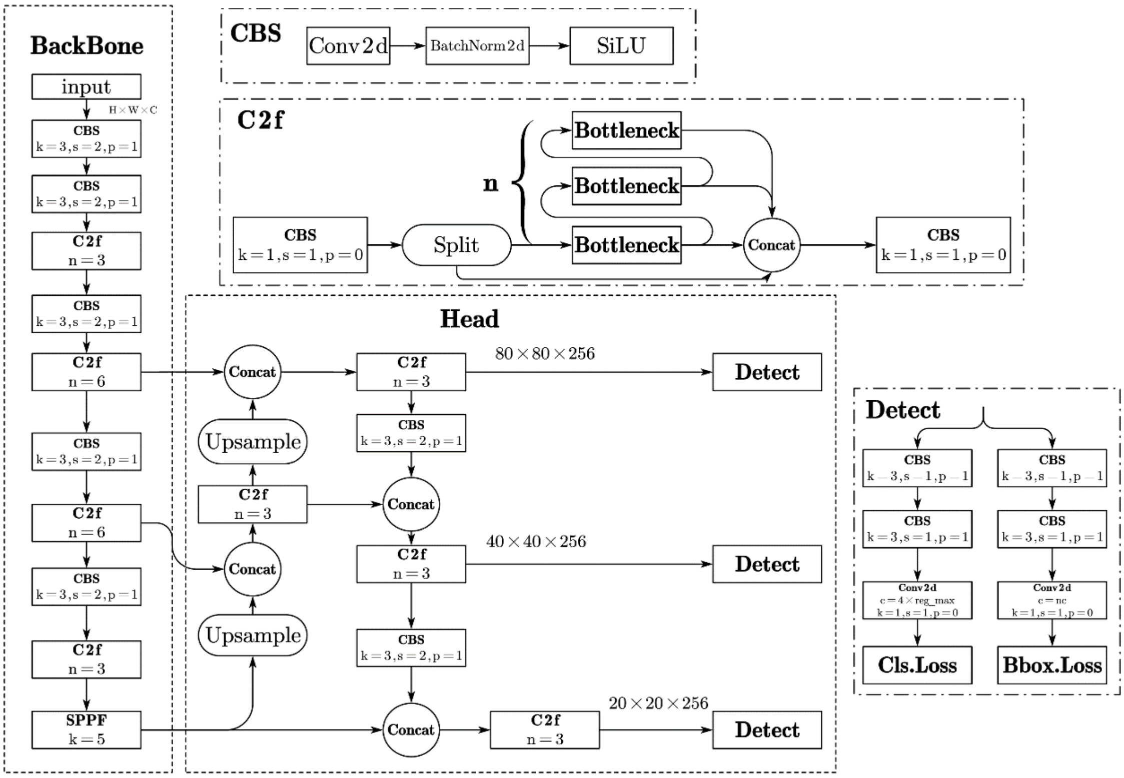

Figure 1. YOLOv8 network structure diagram

3.4. algorithmic process

The target detection model process in this thesis contains the main stages of data processing, feature extraction, feature fusion, target prediction and post-processing. Firstly, data containing targets such as electric vehicles and baby prams are collected through publicly available datasets and surveillance devices and pre-processed, including image size scaling, data enhancement and division into training, validation and test sets. In the data labelling stage, accurate labelling is completed based on categories and target bounding boxes to provide standardised inputs for subsequent model training. The Backbone of the model extracts multi-scale features through convolutional layers and performs feature fusion through Feature Pyramid Network (FPN) and Path Aggregation Network (PAN) to strengthen the multi-scale representation of the target. In the prediction phase, the model generates target bounding boxes, category probabilities and target confidence levels. Subsequently, post-processing removes redundant frames using non-maximal suppression (NMS) technique to ensure the accuracy and reliability of the detection results. Finally, the detection task is completed by outputting the category, bounding box coordinates and confidence score of the target, and the model performance is evaluated by combining the evaluation metrics such as mAP and FPS. The overall framework design is shown in Figure 2.

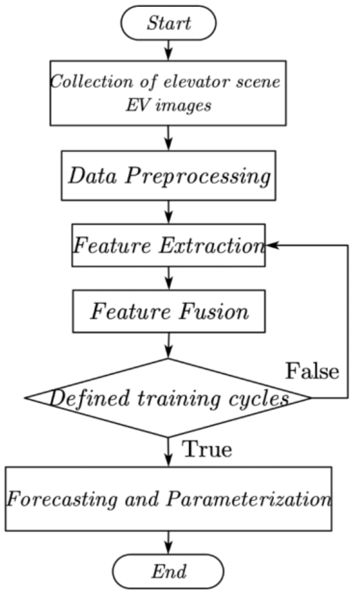

Figure 2. Overall framework flowchart

4. Solution of the model

4.1. Data preprocessing

In this study, a dataset containing multiple target types and scenarios was first collected and organised. The data set mainly consists of target categories such as e-bikes, prams, ordinary bicycles, people and wheelchairs, where e-bikes are used as the main detection objects while other object categories are used as negative samples for reducing false alarms.

During data preprocessing, all images are first converted to JPEG format (.jpg) uniformly to improve storage efficiency and ensure quality. At the same time, the image labels were saved as txt files according to the YOLO format. Since the YOLOv8 model requires the input images to be of the same size, all the images were adjusted to a resolution of 640×640, scaled isometrically and processed with zero-fill to avoid image distortion. In order to effectively improve the robustness of the model, this paper employs a variety of data enhancement methods, including random cropping, horizontal flipping, luminance adjustment and random noise addition, to simulate different scenes and lighting conditions.

The original 6662 images are divided into training, validation and test sets in the ratio of 20:4:1, where 80% is used for training, 10% for validation and 10% for testing. When dividing, it is ensured that the target categories in each subset are evenly distributed to avoid category imbalance affecting the training effect.

4.2. Evaluation indicators

In this study, common target detection evaluation metrics are used to comprehensively assess the performance of the YOLOv8-based electric vehicle in-lift detection system. The main ones include Precision and Recall, which are used to measure the detection accuracy and comprehensiveness of the model. In addition, mean Average Precision Mean (mAP) is used to evaluate the model's comprehensive detection capability on all categories. In order to assess the real-time performance of the model in practical applications, the inference speed (FPS) is also a key metric indicating the processing efficiency of the model.

$$Precision = \frac{TP}{TP+FP} \tag{1}$$

$$\mathrm{Re}call = \frac{TP}{TP + FN} \tag{2}$$

$$mAP = \frac{1}{n}\sum_{i=1}^{n} AP_i \tag{3}$$

Where $Precision$ denotes the proportion of correct predictions made by the model in all samples where the prediction was positive, measuring the ability of the model to misreport. $\mathrm{Re}call$ denotes the proportion of correct predictions made by the model in all samples that were actually positive, measuring the ability of the model to miss a report, TP denotes a positive category that was judged as positive, FP denotes a negative category that was judged as positive, and FN denotes a positive category that was judged as negative. AP_i is the average accuracy of each category, mAP is the average of all categories, which measures the overall detection performance of the model, and n denotes the total number of categories of identified targets.

4.3. Analysis of model output parameters

Monte Carlo tests were conducted to investigate the performance of YOLOv8s, YOLOv5 and MobileNet-SSD in the EV-in-lift scenario with the following results:

YOLOv8s: In multiple simulations, the model's mAP@50 averages 87%, mAP@50-95 is 75%, and the frame rate stays at 125 FPS, which indicates that it can maintain high accuracy and efficiency under complex environments (e.g., light changes, target occlusion).

YOLOv5: Under the same test conditions, the average value of mAP@50 is 85%, and the average value of mAP@50-95 is 74%, with a frame rate of about 120 FPS. Although its accuracy and real-time performance are close to that of YOLOv8s, it is a little bit short of the robustness in complex environments.

MobileNet-SSD: In multiple simulations, the average value of mAP@50 is 81%, the average value of mAP@50-95 is 69%, and the frame rate is about 110 FPS, which is not as good as YOLOv8s and YOLOv5 in terms of high accuracy and real-time performance.

4.4. Analysis of experimental results

Through experimental results and Monte Carlo tests, YOLOv8s shows excellent performance in the EV-in-lift detection task with 87% mAP@50 and 75% mAP@50-95 and achieves a real-time detection speed of 125 FPS on the RTX 4060 GPU. The model maintains good robustness and stability even in complex environments such as light changes and target occlusion. Therefore, YOLOv8s becomes the best choice for the EV-in-lift detection system in this study due to its high accuracy, high real-time performance and strong robustness. The experimental test results are shown in Figure 3, Figure 4, Figure 5, Figure 6.

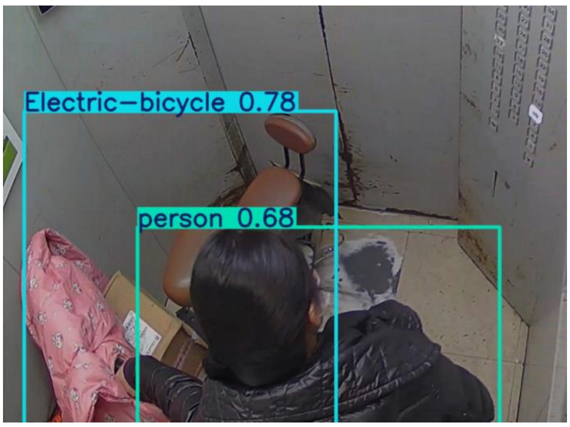

Figure 3. Experimental results 1

Figure 4. Experimental results 2

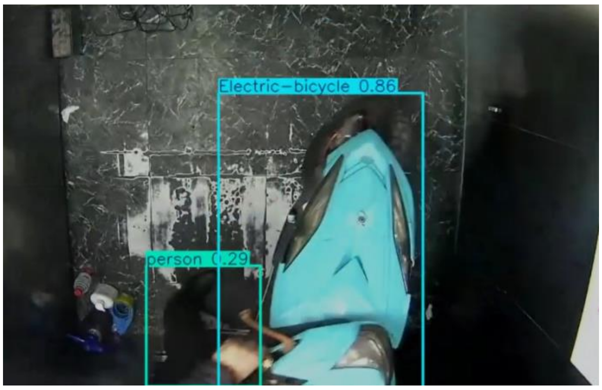

Figure 5. Experimental results 3

Figure 6. Experimental results 4

5. Conclusions

In this study, an efficient electric vehicle into lift detection system is developed based on the YOLOv8s model, which is optimised for practical safety management requirements. By improving the model structure and data processing methods, the system achieves accurate real-time detection of electric vehicles and other targets, with a mAP@50 of 87% and an operating speed of 125 FPS, which meets the deployment requirements of actual surveillance systems. The stable performance in complex environments such as light changes and target occlusion fully demonstrates the application potential of the method in enhancing lift safety management, reducing fire hazards, and improving the overall safety management in urban residential areas.

References

[1] Bigan He. Research and implementation of real-time detection system for electric vehicles in lifts based on improved YOLOv5 [D]. South China University of Technology, 2023.

[2] Yuyang Li, Jiquan Shen, Haixia Zhai, et al. Mask wearing detection algorithm based on improved SSD[J]. Computer Engineering, 2022, 48(08): 173-179+186.

[3] Xin Wang, Yinan Feng. Helmet-wearing detection for cyclists based on improved SSD[J]. Electronic Measurement Technology, 2022, 45(21): 90-97.

[4] Hongcheng Zhao, Ziyang Zhang, Yu Song. [Research on the design of alarm system for electric vehicle intrusion into lift [J]. Fire Protection World (Electronic Edition), 2023, 9 (15): 63-65.

[5] Ziyu Ma. Research on recognition of electric vehicles entering car based on neural network and evidence theory[D]. Zhengzhou University, 2021.

[6] Lin Chen, Chuanjing Chen, Peng Zhang. YOLOv4-tiny based electric vehicle detection model in lifts [J]. Value Engineering, 2022, 41 (12): 117-119.

[7] Bigan He. Lift car electric vehicle banning system based on deep learning and edge computing technology [J]. China Elevator, 2023, 34(11): 15-19+22.

[8] Qiaorun Lu, Hui Lin, Xiaowei Liu. Design of OR-CNN-based dangerous behaviour detection system for electric vehicles entering lifts [J]. Electromechanical Engineering Technology, 2024, 53 (01): 253-256.

[9] Suguang Zhang, Wentao Zhang, Yang Liu, et al. Improved MobileNet-SSD for electric vehicle recognition in lifts [J]. Mechanical Design and Manufacturing, 2024, (09): 340-345.

[10] Yuna Liu, Shuangbao Ma. Lightweight fabric defect detection algorithm based on improved YOLOv8n [J/OL]. Journal of Guangxi Normal University (Natural Science Edition), 1-12 [2024-11-14].

[11] Zhen Liu, Xianzhao Yang, Yang Chen, et al. Target detection algorithm for foggy weather based on improved YOLOv8s [J/OL]. Electronic Measurement Technology, 1-9 [2024-11-14].

[12] Haiqun Wang, Peixu Wei, Haolong Xie, et al. Infrared ship detection based on improved YOLOv8[J/OL]. Electro-Optics and Control, 1-8 [2024-11-14].

[13] Yaoming Ma, Pengfei Zhang, Fusheng Tan. Improved YOLOv8s algorithm for smoke flame detection in complex background [J/OL]., 1-13 [2024-11-14].

[14] Boyuan Wang, Fei Kang, Xuewen Liang, et al. Improved YOLOv8s-based method for identifying surface personnel and seepage in earth and rock dams [J/OL]. 1-13 [2024-11-14].